View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Edinburgh Research Explorer

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Embracing the chaos of behavioral proteomics

Citation for published version:
Bailey, NW & Ly, T 2014, 'Embracing the chaos of behavioral proteomics: a comment on Valcu and Kempenaers' Behavioral Ecology, vol. 26, no. 1, pp. 18-19. DOI: 10.1093/beheco/aru201

Digital Object Identifier (DOI):
10.1093/beheco/aru201

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Behavioral Ecology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. Apr. 2019

Embracing the chaos of behavioral proteomics: a comment on Valcu and Kempenaers

Nathan W. Bailey[a] and Tony Ly[a,b]
[a]Centre for Biological Diversity, School of Biology, University of St Andrews, Greenside Place, St Andrews KY16 9TH, UK and [b]Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, Dow St, Dundee DD1 5EH, UK

It is tempting to feel nihilistic when an avalanche of anonymous omic data is bearing down on you. Particularly if you are studying a trait as variable, flexible, nuanced, and generally misbehaved as behavior. And even more so if you are trying to do it in a nonmodel organism. Nevertheless, Valcu and Kempenaers (2015) argue that behavioral ecologists should take advantage of recent technical advances to apply high-throughput proteomics to behavioral research, much the same way transcriptome profiling now features in many behavioral studies, and they highlight a range of studies that have identified proteins underlying interesting behavioral phenotypes such as sexual conflict and communication. Such progress is encouraging, but we would like to highlight an additional exciting prospect of a behavioral proteomics research program.

We advocate a complementary conceptual approach that takes advantage of the immense complexity and sensitivity of the proteome. High-throughput proteomics is a valuable tool for generating hypotheses about candidate proteins, but as Valcu and Kempenaers (2015) point out, a considerable effort is then required to establish a causal link between protein variation and corresponding behavioral variation, much less the direction of that link. The technical expertise required to reach such a point is likely a much larger hurdle to behavioral ecologists studying nonmodel organisms than is the quantitation of peptides in a mass spectrometry (MS) screen. As an alternative to focused single-protein investigations, it may be useful to directly work with whole-proteome signatures to test hypotheses about the molecular mechanisms and evolutionary origins of interesting behaviors. The unit of study would be the proteome itself, not a protein. This framework is particularly apt for behavioral ecologists; the dynamic range of protein expression can be an order of magnitude greater than that of mRNA expression, proteomes can vary temporally and spatially, and a staggering array of post-translational modifications such as phosphorylation, selective degradation, and differential folding make them exquisitely responsive to environmental perturbation (Ahmad and Lamond 2014).

The chaotic nature of the proteome is one of its most intriguing features, and perhaps the clearest impact of this chaos can be illustrated by considering the proteomic basis of phenotypic plasticity. Understanding how behaviors respond to delicate variations in social or ecological contexts can be enhanced by identifying and categorizing how such environmental noise disturbs the genotype-to-phenotype map. Valcu and Kempanaers (2015) illustrate key studies that describe proteins whose expression correlates with phenotypic plasticity, but these are largely restricted to understanding caste determination in eusocial insects, and they tend to focus on behavioral traits that co-occur with other morphological and developmental changes that may not be readily reversible. In contrast, behaviors such as parental care or differential aggressive responses depending on the social environment (Smiseth and Moore 2002, Logue et al. 2010) are much more dynamic and reversible. Testing proteome-wide patterns associated with such behaviors implicitly acknowledges their complex polygenic basis and environmental sensitivity.

There is precedent for testing hypotheses about whole-proteome variation, and ingenious methods for doing so (Ohta et al. 2010, Khan et al. 2013). Recent work has refined analytical techniques for assessing and testing broad patterns of variation across proteome profiles. For instance, Ly et al. (2014) used global transcriptomics and proteomics analysis to determine the pattern of expression of mRNAs and their cognate proteins across the cell division cycle, and similar approaches could be taken to identify co-ordinated versus discordant mRNA and protein expression levels associated with behavioral phenotypes of interest. Such information would not only clarify the molecular bases of variation in behavior, but could ultimately provide a foundation for testing how selection acting on behavioral variation is—or is not—converted into allele frequency changes. It is important to emphasize the need for explicit hypothesis-testing: for instance, one might test whether less dynamic proteomic components of a phenotype are more resistant to selection, thus channelling evolutionary responses toward more environmentally sensitive proteomic pathways. This could address a longstanding question about the evolution of behavior, which is the relative importance of behavioral flexibility in setting the pace for evolutionary change (West-Eberhard 1989).

We certainly would not argue against candidate gene/protein approaches in behavioral ecology. However, if ever there was a capricious trait likely to be influenced by miniscule, fleeting variations in the expression of a large number of proteins, it is behavior. An additional advantage of a systems approach is that testing hypotheses about behavioral proteomics need not rely on highly detailed gene annotations (Wühr et al. 2014). For example, experimental evolution studies such as are performed in *Drosophila* lines subjected to varying opportunity for sexual selection could assess whether more socially responsive constituents of the genotype-to-phenotype map show correspondingly slower or faster responses to selection, without needing to know the function of the genes involved (Immonen et al. 2014). These genotype-to-phenotype maps are already being used to enrich the annotation of the human genome by associating molecular signatures to both simple and complex phenotypes, such as gene deletion and disease (Subramanian et al. 2005). In the immediate future, we anticipate the "cleverest" experiments will use the avalanches of anonymous "omic data currently being generated to deliver insights into the evolutionary and molecular constraints—and evolutionary and molecular paths-of-least-resistance—that cause interesting behavioural variation in nature. Embrace the chaos of the proteome!

Address correspondence to N.W. Bailey. E-mail: nwb3@st-andrews.ac.uk.

Received 9 October 2014; accepted 11 October 2014; Advance Access publication 11 November 2014.

doi:10.1093/beheco/aru201

Editor-in-Chief: Leigh Simmons

REFERENCES

Ahmad Y, Lamond AI. 2014. A perspective on proteomics in cell biology. Trends Cell Biol. 24:257–264.

Immonen E, Snook RR, Ritchie MG. 2014 Mating system variation drives rapid evolution of the female transcriptome in *Drosophila pseudoobscura*. Ecol Evol. 4:2186–2201.

Khan Z, Ford MJ, Cusanovich DA, Mitrano A, Pritchard JK, Gilad Y. 2013. Primate transcript and protein expression levels evolve under compensatory selection pressures. Science. 342:1100–1104.

Downloaded from https://academic.oup.com/beheco/article-abstract/26/1/18/2263121
by guest
on 14 March 2018

Logue DM, Abiola IO, Rains D, Bailey NW, Zuk M, Cade WH. 2010. Does signalling mitigate the cost of agonistic interactions? A test in a cricket that has lost its song. Proc Biol Sci. 277:2571–2575.
Ly T, Ahmad Y, Shlien A, Soroka D, Mills A, Emanuele MJ, Stratton MR, Lamond AI. 2014. A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. Elife. 3:e01630.
Ohta S, Bukowski-Wills JC, Sanchez-Pulido L, Alves Fde L, Wood L, Chen ZA, Platani M, Fischer L, Hudson DF, Ponting CP, et al. 2010. The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. Cell. 142:810–821.
Smiseth PT, Moore AJ. 2002. Does resource availability affect offspring begging andparental provisioning in a partially begging species? Anim Behav. 63:577–585.
Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 102:15545–15550.
Valcu C-M, Kempenaers B. 2015. Proteomics in behavioral ecology. Behav Ecol. 26:1–15.
West-Eberhard MJ. 1989. Phenotypic plasticity and the origins of diversity. Ann Rev Ecol Syst. 20:249–278.
Wühr M, Freeman RM Jr, Presler M, Horb ME, Peshkin L, Gygi SP, Kirschner MW. 2014. Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. Curr Biol. 24:1467–1475.

Finding one's way through the proteome: a response to comments on Valcu and Kempenaers

Cristina-Maria Valcu and Bart Kempenaers
Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard-Gwinner-Street 8, D-82319 Seewiesen, Germany

It is encouraging to see that the authors of the commentaries to our review share our enthusiasm regarding the promising perspectives of applying proteomics to behavioral ecology. We have explained how the proteomic approach can assist behavioral ecologists in understanding the molecular basis of behavior, as well as variation in and evolution of behavior (Valcu and Kempenaers 2015). The 3 commentaries provide additional arguments and examples supporting this view (Bailey and Ly 2014; Ramm 2014; Sirot 2014). Here, we would like to comment on an interesting topic raised by Bailey and Ly (2014).

In their commentary, Bailey and Ly draw attention to a particular approach to proteomic data analysis, namely data mining. They advocate the use of whole-proteome signatures for testing hypotheses about behavior. Proteomic signatures (aka protein expression patterns in our review, protein expression signatures, Bradley et al. 2002; protein expression profiles, Shen et al. 2013; proteomic signature profiles, Goh et al. 2012; proteomic profiles or patterns, Petricoin et al. 2002) represent patterns of protein abundance, which are indicators of particular phenotypes or biological conditions. The interpretation of these patterns does not require further information about the identity of the proteins. Proteins changing in abundance collectively contribute to the patterns, irrespective of what caused the change (e.g., gene expression up- or downregulation, protein turnover, or modification). Hence, proteomic signatures are more powerful in discriminating phenotypes than variation in any of the single proteins they comprise. This is useful particularly for heterogeneous populations (Petricoin et al. 2002) and for highly variable phenotypes such as behavior.

Proteomic signatures can be obtained through an independent selection of differentially expressed proteins based on statistical criteria or can be extracted from complex proteomic data sets using machine learning algorithms, such as those suggested by Bailey and Ly (2014). In the latter case, the choice for the pattern recognition algorithm depends on the data available and on the desired output. For example, supervised learning can be employed for pattern recognition in data with an already known structure (e.g., treatment vs. control), whereas unsupervised learning assists the discovery of previously unknown patterns without making assumptions about a structure in the data (Thomas et al. 2006).

Proteomic signatures have been long recognized as useful tools with applications, for example, in diagnostic and disease monitoring (Petricoin et al. 2002), pharmacology (Wenzel and Bandow 2011), toxicology (Amacher 2010), ecotoxicology (Tomanek 2011), and ecology (Renella et al. 2014). Such global proteomic signatures identified based on either protein presence/absence (Biron et al. 2005; Ponton et al. 2006; Lefèvre et al. 2007) or protein abundance (Chan et al. 2011) have also been used in some of the behavioral ecology studies we reviewed. Data mining is a powerful approach to identify hidden phenotypes because it uses proteome-wide information on protein presence or abundance, not only subsets of proteins that satisfy certain criteria (e.g., differentially expressed). This can, for example, help revealing groups of individuals with diverging molecular phenotypes within otherwise (behaviorally) homogenous groups. As Bailey and Ly (2014) also point out, whole-proteome signatures encompass many small differences in protein abundances scattered across the proteome, and this makes them a sensitive tool for investigating the molecular basis of variation in behavior and the evolution of behavior. Furthermore, whole-proteome signatures allow tackling phenomic studies (i.e., genotype-to-phenotype mapping) (Houle et al. 2010; Bailey and Ly 2014).

We feel, however, that a note of caution is needed here. The results of heuristic algorithms largely depend on the data being analyzed (Thomas et al. 2006) and computer scientists warn that "data mining is easy to do badly" (Larose 2014). The solutions identified may not be unique and require extensive validation (Thomas et al. 2006). From a technical perspective, gel-based approaches may suffer from incomplete separation of proteins (as discussed in our review), which makes them less suitable for data mining approaches because 1 band or spot often contains more than 1 protein. These limitations probably explain the tendency of proteomic studies to favor traditional statistic tools for data analysis. However, when carefully used, bioinformatic tools typically employed for the analysis and interpretation of proteomic data should produce consistent results whether applied on preselected protein sets or on whole-proteome data sets (Huang et al. 2009).

On the other hand, although proteomic signatures in the absence of protein identity are undoubtedly valuable tools for data analysis, they only become truly insightful when incorporating prior knowledge on protein function (Subramanian et al. 2005). We strongly believe that the full potential of proteomic tools in helping us to understand the molecular basis of behavior will only be reached by learning the identity and the function of the proteins comprising a behavior-specific proteomic signature.

Whatever the approach undertaken for analysis and however challenging high-throughput proteomic studies may be, of one thing we can be sure: the proteome holds answers to many of the questions asked by behavioral ecologists and searching for them will be worth the effort!

Address correspondence to C.-M. Valcu. E-mail: cvalcu@orn.mpg.de.

Received 17 November 2014; accepted 19 November 2014; Advance Access publication 22 December 2014.

doi:10.1093/beheco/aru225

Editor-in-Chief: Leigh Simmons

Downloaded from https://academic.oup.com/beheco/article-abstract/26/1/18/2263121
by guest
on 14 March 2018